DEPARTMENT OF APPLIED ECONOMICS

RESEARCH REPORT

DECOMPOSING GRANGER CAUSALITY OVER THE SPECTRUM

AURÉLIE LEMMENS • CHRISTOPHE CROUX • MARNIK G. DEKIMPE

OR 0454

KATHOLIEKE UNIVERSITEIT LEUVEN

Decomposing Granger Causality over the Spectrum

Aurélie Lemmens[1] †, Christophe Croux† and Marnik G. Dekimpe†‡

†Department of Applied Economics, K.U.Leuven

‡Faculty of Business Administration, Erasmus University Rotterdam

November 16, 2004

[1]Corresponding author. Phone: +32-16-326960; Fax: +32-16-326732. Naamsestraat 69, B-3000 Leuven, Belgium. Email: aurelie.lemmens@econ.kuleuven.ac.be. This research has been funded by the Research Fund K.U.Leuven and the "Fonds voor Wetenschappelijk Onderzoek" under contract numbers G.0385.03, G.0594.05 and G.0116.04.

JEL Classification: C22 , C53, E23.

Abstract

We develop a bivariate spectral Granger-causality test that can be applied at each individual frequency of the spectrum. The spectral approach to Granger causality has the distinct advantage that it allows to disentangle (potentially) different Granger-causality relationships over different time horizons. We illustrate the usefulness of the proposed approach in the context of the predictive value of European production expectation surveys.

KEYWORDS: Business Surveys, Granger causality, Production Expectations, Spectral Analysis.

1 Introduction

Investigating causality is a topic of main interest in scientific research. To assess the causality between two processes in a common and well-defined (non-experimental) framework, one usually refers to the well-known concept of *Granger* causality (GC), introduced in 1969 by the 2003 Nobel prize winner in Economics. GC reflects a restricted sense of causality, i.e. *the extent to which a process X_t is leading another process Y_t*, and builds upon the notion of incremental predictability. Specifically, a process X_t Granger causes another process Y_t if future values of Y_t can be better predicted using the past values of X_t and Y_t rather than only past values of Y_t. The reader should keep in mind that, in some circumstances, the aforementioned notion of causality may not fully coincide with the concept of causation (Zellner, 1979). However, the underlying intuition for this approach is that if an event is the cause of another, it should precede it. Therefore, as soon as the application of the causality concept refers to the search for the best predictive model, the concept of GC is valid. The standard test of GC developed by Granger (1969) is based on the following regression model

$$Y_t = \alpha_0 + \sum_{k=1}^{M} \beta_{1k} Y_{t-k} + \sum_{k=1}^{M} \beta_{2k} X_{t-k} + \varepsilon_t, \qquad t = M+1, ..., T \tag{1}$$

where ε_t are uncorrelated random variables with mean zero and variance σ^2, and M is the specified lag length. The null hypothesis that X_t does not Granger cause Y_t is supported when $\beta_{2k} = 0$ for $k = 1, ..., M$, causing (1) to reduce to

$$Y_t = \alpha_0 + \sum_{k=1}^{M} \beta_{1k} Y_{t-k} + \tilde{\varepsilon}_t. \tag{2}$$

A wide range of bivariate GC tests exist,[1] which have been used extensively to study a wide range of substantive, economic, issues. For instance, the well-known "export-led growth" hypothesis has been studied repeatedly in a GC framework (see

[1]Two well-known alternative test statistics can be derived from equations (1) and (2), i.e. the

e.g. Abual-Foul, 2004), as has been the relationship between economic growth and various other variables, such as business-cycle volatility (Dopke, 2004), the degree of openness (Bahmani-Oskooee and Niroomand, 1999), and defense spending (Al-Yousif, 2002). The causal relationship between money and output has also been studied extensively (Cheung and Fujii, 2001). In the financial literature, GC testing has been applied, for instance, to identify price-leadership patterns among national stock prices (Peiers, 1997), to study the stock price-volume relationship (Hemstra and Jones, 1994), to get insight in the dynamic behavior of bonds and stocks (Park and Shenoy, 2002), or in the international links between interest rates (Bruneau and Jondeau, 1999). In marketing, GC testing has been used predominantly to discern competitive reactions patterns (see Hanssens *et al.*, 2001, p.314 for a review). While not exhaustive, the above enumeration clearly demonstrates the widespread use of the GC concept in both economics and business.

In this paper, we propose a spectral-density based GC test. This approach offers

Granger-Sargent and the Granger-Wald test (1969). The Granger-Sargent test is defined by

$$GS = \frac{(R_2 - R_1)/M}{R_1/(T - 2M)}$$

where R_1 is the residual sum of squares in model (1) and R_2 is the residual sum of squares in model (2). The Granger-Sargent test statistic has an F-distribution with M and $(T - 2M)$ degrees of freedom. The Granger-Wald test, in turn, is defined as

$$GW = T\frac{(\hat{\sigma}^2_{\tilde{\varepsilon}_t} - \hat{\sigma}^2_{\varepsilon_t})}{\hat{\sigma}^2_{\varepsilon_t}}$$

where $\hat{\sigma}^2_{\tilde{\varepsilon}_t}$ is an estimate of the variance of $\tilde{\varepsilon}_t$ from model (2), and $\hat{\sigma}^2_{\varepsilon_t}$ is an estimate of the variance of ε_t from model (1), and follows an asymptotic χ^2_M distribution under the null hypothesis. Both tests are asymptotically equivalent. However, for smaller sample sizes, the Granger-Wald test has been shown to have more power than the Granger-Sargent test, and is therefore better suited for identification purposes. A third common GC test is the double prewhitening technique known as the Haugh-Pierce (1977) test (See Section 2 for more details). This test has been found to have a reasonable power together with a small bias in size (type-I error) (see e.g. Bult *et al.*, 1997 for a review).

some distinct advantages relative to the aforementioned, standard, procedures. While traditional tests indicate whether Granger causality is present or not, we propose to give a richer and more complete picture by decomposing Granger causality over different time horizons. As such, one can, for example, compare the predictive power present at the short, middle or long run. The spectral GC test can be applied at any given frequency of the spectrum. This allows us to gain insights into potential variations in the strength of the GC between the two variables over the spectrum. Indeed, there is increasing research evidence that the nature of the relationship between two variables may vary depending on the time horizon under consideration. Such variation was, for example, found in the relationship between real exchange rates and real interest differentials (Baxter, 1994), between the GDP series of different countries (Croux *et al.*, 2001), in the nature of competitive price reactions (Bronnenberg *et al.*, 2004), and in the link between aggregate advertising spending and various macro-economic indicators (Deleersnyder *et al.*, 2003). Baxter (1994), for example, found evidence of a relationship between real exchange rates and real interest differentials at trend (long-run) and business-cycle (middle-run) frequencies (i.e. low to middle frequencies), which was not found in prior studies that only focused on high-frequency components. Similarly, Croux *et al.* (2001) found that the GDPs of US states are more correlated with each other than European countries' together, but also that this difference is much more pronounced in the short run (i.e. at the high frequencies). As a consequence, a one-shot GC test that is supposed to apply across all time horizons (e.g. in the short run, over the business cycle frequencies, and in the long run) may well give an incomplete, and potentially misleading, picture of the temporal ordering between the variables of interest.

An important feature related to spectral-based GC tests is that they are asymptotically independent of each other at different frequencies. This allows us to carry out a joint test for GC at different frequencies of the spectrum simultaneously. Note that estimated cross-correlations between X_t and Y_t are, in contrast, strongly correlated. Moreover, cross-correlations are very cumbersome to interpret, since there is confoundedness between correlations *within* the series and *between* the series (e.g. Chatfield, p. 139, 1996). Hence, the cross-correlogram between Y_t and X_t is not an appropriate tool to study the decomposition of the GC over different time horizons. We will rather use the value of the spectral-based GC test as a measure of GC and plot these measures with respect to the frequency, resulting in a graphical tool which gives insight into the decomposition of the GC over the spectrum. This plot is essentially given by the Fourier transform of the cross-correlations at negative lags between filtered versions of the series X_t and Y_t.

The remainder of paper is structured as follows. The spectral GC test is detailed in Section 2. In Section 3, we illustrate its use in studying the predictive value of European production expectation surveys. Section 4 contains a brief conclusion.

2 A Spectral Granger-Causality Approach

Let X_t and Y_t be stationary (after possible transformations) time series. Spectral analysis is performed on the innovations series, u_t and v_t, derived from X_t and Y_t. The latter are modelled as univariate ARMA processes, i.e.

$$\begin{aligned} \Phi^x(L)X_t &= C^x + \Theta^x(L)u_t \qquad (3) \\ \Phi^y(L)Y_t &= C^y + \Theta^y(L)v_t \end{aligned}$$

where $\Phi^x(L)$ and $\Phi^y(L)$ are autoregressive polynomials, $\Theta^x(L)$ and $\Theta^y(L)$ moving-average polynomials, and C^x and C^y potential deterministic components. After filtering the series with the above ARMA models, we obtain the innovation series u_t and v_t, which are white-noise processes with zero mean, possibly correlated with each other at different leads and lags. These innovations are central to the development of the well-known Haugh-Pierce (1977) test for GC,[2] and will also form the main building blocks for our proposed testing procedure.

Let $S_u(\lambda)$ and $S_v(\lambda)$ be the spectral density functions, or spectra, of u_t and v_t at frequency $\lambda \,\epsilon\, [-\pi, \pi]$ defined by

$$S_u(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_u(k) e^{-i\lambda k} \quad \text{and} \quad S_v(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_v(k) e^{-i\lambda k}, \tag{4}$$

where $\gamma_u(k) = Cov(u_t, u_{t-k})$ and $\gamma_v(k) = Cov(v_t, v_{t-k})$ represents the autocovariances of u_t and v_t at lag k. The idea of the spectral representation is that each time series may be decomposed into a sum (or integral) of uncorrelated components, each related to a particular frequency λ. A detailed treatment on the spectral analysis of time series is given in Koopmans (1995).

As the innovations series u_t and v_t are white-noise processes, the spectra (4) are constant functions, given by

$$S_u(\lambda) = \frac{Var(u_t)}{2\pi} \quad \text{and} \quad S_v(\lambda) = \frac{Var(v_t)}{2\pi}.$$

Therefore, their spectra $S_u(\lambda)$ and $S_u(\lambda)$ can simply be estimated as

$$\hat{S}_u(\lambda) = \frac{\widehat{Var}(u_t)}{2\pi} \quad \text{and} \quad \hat{S}v(\lambda) = \frac{\widehat{Var}(v_t)}{2\pi}.$$

To investigate the relationship between both stochastic processes under consideration, we consider the cross-spectrum, $S_{uv}(\lambda)$, between u_t and v_t. This is a complex

[2]Under the null hypothesis of no GC, the M cross-correlations $\rho_{vu}(k) = corr(v_t, u_{t-k})$ between the innovation series u_t and v_t, with $k = 1, ..., M$, are asymptotically independently and normally distributed with mean zero and standard deviation $T^{-1/2}$, and the Haugh-Pierce (1977) test statistic

$$HP = T \sum_{k=1}^{M} \hat{\rho}^2_{vu}(k)$$

is asymptotically chi-square distributed with M degrees of freedom.

number, defined as

$$S_{uv}(\lambda) = C_{uv}(\lambda) + iQ_{uv}(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{uv}(k) e^{-i\lambda k}, \tag{5}$$

where the cospectrum $C_{uv}(\lambda)$ and the quadrature spectrum $Q_{uv}(\lambda)$ are, respectively, the real part and the imaginary part of the cross-spectrum. Here $\gamma_{uv}(k) = Cov(u_t, v_{t-k})$ represents the cross-covariance of u_t and v_t at lag k. The cross-spectrum can be estimated non-parametrically by

$$\hat{S}_{uv}(\lambda) = \frac{1}{2\pi} \left\{ \sum_{k=-M}^{M} w_k \hat{\gamma}_{uv}(k) e^{-i\lambda k} \right\}, \tag{6}$$

with $\hat{\gamma}_{uv}(k) = \widehat{Cov}(u_t, v_{t-k})$, the empirical cross-covariances, and with window weights w_k, for $k = -M, \dots, M$. The expression (6) is called the *weighted covariance estimator*, and when the weights w_k are selected as $1 - \frac{|k|}{M}$, the Barlett weighting scheme is obtained. The constant M is the maximum lag order considered.[3]

This cross-spectrum allows to compute the coefficient of coherence $h_{uv}(\lambda)$, defined as (see Koopmans, 1995)

$$h_{uv}(\lambda) = \frac{|S_{uv}(\lambda)|}{\sqrt{S_u(\lambda) S_v(\lambda)}}. \tag{7}$$

This coefficient, which can take on values between zero and one, gives a symmetric measure of the strength of linear association between two time series, frequency by frequency, but does not express any information on the *direction* of the relationship between two processes. The squared coefficient of coherence, which has a similar interpretation as the R-squared in a regression context, was used in Barksdale *et al.* (1974) to study the association between advertising and sales, or in Woitek (2003) to investigate the relationship between human height cycles and cycles of economic variables.

[3] In practice (see e.g. Diebold, 2001, p.136), M is often chosen to be equal to the square root of the number of observations T.

A confidence interval for the coefficient of coherence can be derived. Specifically, for $n > 20$, we have that $\tanh^{-1}(h_{uv}(\lambda))$ is well approximated by a normal distribution:

$$\tanh^{-1}\left(\hat{h}_{uv}(\lambda)\right) \approx N\left(\tanh^{-1}(h_{uv}(\lambda)), \frac{1}{2(n-1)}\right), \tag{8}$$

with $n = T/\sum w_k^2$ (see e.g. Koopmans, 1995). We denote $2(n-1)$ as the equivalent degrees of freedom, EDF. From the above equation, if follows immediately that one can reject the null hypothesis of the nullity of the coefficient of coherence if

$$|\hat{h}_{uv}(\lambda)| > \tanh\left(z_{\alpha/2}\sqrt{\frac{1}{EDF}}\right),$$

where $z_{\alpha/2}$ is the $\frac{\alpha}{2}$ upper quantile of a standard normal distribution.

If one wants to take the *direction* of the relationship into account, as asked for to indicate GC, the coefficient of coherence in (7) should be adapted. Specifically, the cross-spectrum (5) can be decomposed into three parts, (i) $S_{u\Leftrightarrow v}$, the instantaneous relationship between u_t and v_t, (ii) $S_{u\Rightarrow v}$, the directional relationship between v_t and lagged values of u_t, and (iii) $S_{v\Rightarrow u}$, the directional relationship between u_t and lagged values of v_t, i.e.

$$\begin{aligned} S_{uv}(\lambda) &= \frac{1}{2\pi}\left[S_{u\Leftrightarrow v} + S_{u\Rightarrow v} + S_{v\Rightarrow u}\right] \\ &= \frac{1}{2\pi}\left[\gamma_{uv}(0) + \sum_{k=-\infty}^{-1}\gamma_{uv}(k)e^{-i\lambda k} + \sum_{k=1}^{\infty}\gamma_{uv}(k)e^{-i\lambda k}\right]. \end{aligned} \tag{9}$$

The proposed spectral measure of GC is based on the key property that X_t does not Granger cause Y_t if and only if $\gamma_{uv}(k) = 0$ for all $k < 0$ (Gouriéroux *et al.*, 1996). Hence, if the goal is to assess the predictive content of X_t relative to Y_t, one is mainly interested in the second part of (9), i.e.

$$S_{u\Rightarrow v}(\lambda) = \frac{1}{2\pi}\left[\sum_{k=-\infty}^{-1}\gamma_{uv}(k)e^{-i\lambda k}\right].$$

A Granger coefficient of coherence is then given by

$$h_{u\Rightarrow v}(\lambda) = \frac{|S_{u\Rightarrow v}(\lambda)|}{\sqrt{S_u(\lambda)S_v(\lambda)}}. \tag{10}$$

Therefore, in the absence of GC, given that the numerator in 10 cancels out, $h_{u\Rightarrow v}(\lambda) = 0$ for every λ in $[-\pi, \pi]$. A natural estimator for the Granger coefficient of coherence at frequency λ is

$$\hat{h}_{u\Rightarrow v}(\lambda) = \frac{\left|\hat{S}_{u\Rightarrow v}(\lambda)\right|}{\sqrt{\hat{S}_u(\lambda)\hat{S}_v(\lambda)}},$$

with $\hat{S}_{u\Rightarrow v}(\lambda)$ as in (6), but with all weights w_k for $k \geq 0$ put equal to zero. Similarly as for the coefficient of coherence distribution (8), one has that the transformed GC coefficient is approximately normally distributed[4]

$$\tanh^{-1}\left(\hat{h}_{u\Rightarrow v}(\lambda)\right) \approx N\left(\tanh^{-1}\left(h_{u\Rightarrow v}(\lambda)\right), \frac{1}{2(n'-1)}\right) \tag{11}$$

with

$$n' = \frac{T}{\sum_{k=-M}^{-1} w_k^2}.$$

Indeed, since the weights w_k with a positive index k are set equal to zero when computing $\hat{S}_{u\Rightarrow v}(\lambda)$, only the w_k with negative indices need to be taken into account when computing the appropriate degrees of freedom $EDF' = 2(n'-1)$. The null hypothesis of no Granger causality at frequency λ, formally $H_0 : h_{u\Rightarrow v}(\lambda) = 0$, is then rejected if

$$|\hat{h}_{u\Rightarrow v}(\lambda)| > \tanh\left(z_{\alpha/2}\sqrt{\frac{1}{EDF'}}\right). \tag{12}$$

Another type of decomposition of the cross-spectrum between two time series was proposed by Gouriéroux *et al.* (p.373-382, 1996). Our approach differs from theirs in a number of ways. First, and most importantly, we derive critical values accompanying the GC measures. This allows for more formal statistical inference than previously possible. Second, unlike the non-parametric estimation of the spectra proposed in

[4]The weights in (6) are usually taken to be symmetric, in the sense that $w_{-k} = w_k$. Allowing for asymmetric weights, however, does not alter the asymptotic results, as can be seen by verifying the proofs in Brillinger (1969).

this paper, spectra in Gouriéroux *et al.* need to be estimated parametrically, via a bivariate Vector Autoregressive model. This works well for time series which are relatively smooth. However, the parametric estimation is rather insensitive to local variations of the spectrum. Consequently, it may overlook spectral peaks (Koopmans, 1995).

Our spectral-based GC approach is very flexible in the sense that it provides a measure of GC at each individual frequency λ of choice, making it feasible to investigate the strength of the causal relationships, e.g. at the short run, at the business cycle frequency, or at the long run. Relative to the other existing GC tests (cf. footnote 1), the spectral-based approach therefore provides additional information on the time horizon where GC is (most) prevalent, or, in contrast, negligible.

Finally, we want to point out that it is easy to perform a joint test for a null hypothesis of the form $H_0 : h_{u\Rightarrow v}(\lambda_1) = \ldots = h_{u\Rightarrow v}(\lambda_s) = 0$, with $\lambda_1, \ldots, \lambda_s$ being different frequencies. Indeed, under H_0, one has

$$EDF \sum_{j=1}^{s} \left(\tanh^{-1}(\hat{h}_{u\Rightarrow v}(\lambda_j))\right)^2 \approx \chi_s^2, \tag{13}$$

since the spectral estimates $\hat{h}_{u\Rightarrow v}(\lambda)$ are asymptotically independent of each other at different frequencies. In the following application, we illustrate the use of such a joint test.

3 Application

3.1 Introduction

Various governments conduct, at regular intervals, a wide range of surveys about the judgments and anticipations of consumers, producers and/or manufacturers. As one of the largest-scale surveys, the European Union has collected, for over forty years, so-called Business Tendency Surveys. About 68,000 companies and 27,000 consumers across the European Union are surveyed each month about (i) their *judgments* (i.e. their assessment of the current or past status of a given variable), and (ii) their *expectations* (i.e. their estimation of the likely future status of that variable). Even though such surveys are costly to conduct, they may offer useful leading information on the underlying economic variables. Indeed, the actual, objectively measured, values of those variables, denoted as *accounts*, typically become available only several months later than the surveys' results (Buffeteau and Mora, 2000). However, timeliness may not be a sufficient condition to justify the high costs involved, and the accuracy of the Business Tendency Surveys should be another key consideration. Specifically, one should expect (or hope) that these attitudinal measures have good predictive power, a question that naturally translates into the notion of GC. If a surveyed variable Granger causes its complementary account variable, the respondents of the survey would be found to possess implicit knowledge about the future levels of the account variable, that could not be derived from previous account levels.

Even though various studies have already considered whether judgmental data reported in the European Union's Business Tendency Surveys Granger cause the corresponding objective measures (see, in this respect, Hanssens and Vanden Abeele, 1987 or Lemmens *et al.*, 2004), none of them has decomposed this potential predictive power over different time horizons. Using the procedure outlined in Section 2, it is now possible to compare the strength of the GC for the short-run, middle-run and longer-run components of the series. We can test, for example, whether the strength of the causality at the short run is significant or not. If the former is verified, the Business Tendency Surveys are able to pick up the quickly-changing (i.e. short-run) components of the series. Our prior belief is that the slowly-moving (i.e. long-run) component would be more accurately forecasted by the surveys' respondents than the more unpredictable fast-moving changes in production. Testing whether the GC is significant at a long run frequency boils down to check whether the surveys have predictive value for the more slowly-varying component of the corresponding account series.

3.2 Data

In this illustration, we focus on the (potential) predictive value of the European Production Expectations series. These publicly available data[5] are provided by the Directorate General Economy and Finance of the European Union. They reflect the respondents' optimism/pessimism w.r.t. the evolution of the production, and are expressed in Balance ($Bal = Pos - Neg$). Specifically, one asks the responding firms whether they expect certain variables to increase, decrease or remain stable over time, and subsequently subtracts all *decrease* (Neg) answers - in percentage points of total answers - from the percentage of *increase* (Pos) ones. This expectation series will form our X_{it} series (with $i = 1, ..., 12$), as data will be used on 12 countries (i.e. Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, The Netherlands and the United Kingdom).[6] The data range from January 1985 to December 2002, resulting in 216 observations.

The investigated account time series, i.e. the Y_{it} series for each country i, are the

[5]See http://europa.eu.int/ comm/economy_finance/ indicators/businessandconsumersurveys_en.htm

[6]Three countries (Portugal, Spain and Sweden) were not withheld, as this would have resulted in the loss of multiple data points due to missing observations, since surveys in these countries began later.

European Production Accounts series, which are published as part of the National Accounts Statistics by the OECD.[7] They are expressed as an index with 1995 scaled as base index 1 (at constant prices). All time series are collected on a monthly basis, and are already seasonally adjusted by the data providers. In the sequel, we formally investigate whether the European production expectations have significant predictive power, and if so, at what time frame, with respect to their corresponding future production accounts.[8]

3.3 Empirical Results

Since the traditional time-domain-based GC tests as well as the proposed spectral GC test require stationarity of time series, we seasonally differenced the production account series[9] (Y_{it}), as in Lemmens *et al.* (2004). In line with prior studies (see e.g. Hanssens and Vanden Abeele, 1987; Öller and Tallbom, 1996), we found that all production expectation series (X_{it}) were already stationary. Next, the series were filtered[10] to obtain white-noise processes, which, as indicated in Section 2, form the building blocks for our Granger coefficient of coherence.

For each of the twelve countries, we present in Figure 1 the estimated Granger coefficients of coherence $\hat{h}_{u\Rightarrow v}(\lambda)$, for all $\lambda \in [0, \pi]$. This coefficient tests whether the production expectations are Granger causing the production accounts of a given country at that frequency. The baseline represents the critical bound of significance at the 5% probability level, as given in (12). Note that the frequency λ on the horizontal axis can be translated into a cycle of T months by $T = 2\pi/\lambda$ (in months, for monthly data).

[7]See OECD publication, *Main Economics Indicators*, the Industrial Production Index, ref. 2027K

[8]For all undergone tests, we took $M = \sqrt{T}$

[9]The Irish account series was taken in logarithm due to its exponential trend

[10]According to different diagnostic tests, residuals obtained after *SARIMA* modeling did not deviate significantly from white-noise processes (Figures are available upon request).

[Insert Figure 1 about here]

Figure 1 shows a consistent pattern across eight countries, i.e. Austria, Belgium Finland, France, Germany, Italy, The Netherlands, and the United Kingdom. For those countries, the GC at small frequencies (corresponding to the longer-run components) clearly dominates those at the higher frequencies (corresponding to the shorter-run components). Hence, even though these countries' production expectations are found to have significant (incremental) predictive power with respect to the longer-run components in the production account series, they have much more difficulty in predicting the fast-moving components of these series. Apart from Germany, the Granger coefficients of coherence corresponding with the high frequencies hardly reach statistical significance, and also for Germany, the GC measure remains much more pronounced at the lower frequencies. However, given these countries' significant Granger coefficient of coherence at the lower part of the frequency band, we expect, for these countries, that also an overall GC test will indicate significance. As indicated in Table 1, this conjecture was confirmed through a formal Granger-Wald test in seven instances ($p < 0.01$), while the test statistic for the remaining country (The Netherlands) was only marginally insignificant ($p = 0.056$). A similar picture emerged when applying the Haugh-Pierce test as "overall" GC test, even though the corresponding p-values were somewhat higher.[11] A possible explanation for this loss in power is that for these eight countries, the GC is concentrated with the low-frequency components (see Figure 1), while the Haugh-Pierce test does not give enough weight to this part of the spectrum.

Other countries, in contrast, do not show much evidence of GC between production expectations and accounts at any of the frequencies. This is the case for Greece, Ireland, and, to a lesser extent, Denmark and Luxembourg. As a consequence, for those countries, we do not expect an overall GC test to give a significant outcome

[11] This affected the outcome of the test statistic in only one instance, i.e. the United Kingdom.

either. Again using the conventional Granger-Wald test, we found support for this conjecture for Ireland ($p = 0.357$) and Denmark ($p = 0.146$). For Luxembourg, a significant overall test statistic was found ($p < 0.05$), which may be attributed to the significant GC coefficients of coherence between frequencies 1.5 and 2.5. This significant overall GC corresponds, however, with a very different coherence pattern than the one observed for the previous eight countries. This difference in the nature of the GC relationship over different time horizons goes undetected in conventional test procedures. When applying the Haugh-Pierce test, an insignificant effect was found for three countries (Denmark, Greece and Ireland), while, as with the Granger-Wald test, a significant overall GC was found for Luxembourg.

A few countries also have some idiosyncratic features. For example, we observe a small increase in the GC coherence measure around frequency 1.5 (corresponding with cycles of 4 months) in Belgium, France and Italy, while Finland experiences such a small increase around frequency 1.0. Rather than trying to explain each of these idiosyncratic features, we find it more insightful to focus on the general picture that emerges across the various countries. Specifically, the Granger coefficient of coherence tends to either decrease in λ (for eight out of twelve countries), or to remain flat and non-significant over the whole frequency band (see e.g. Greece, Denmark and Ireland). For the latter countries, the production expectations do not convey additional information about future production levels, which puts into question the usefulness of spending considerable amounts of money in collecting these data. Such additional information is conveyed for the other countries, however, but (except for Germany) *only* for the longer-run evolution of the production series. Hence, while production expectation series may be available earlier than the actual account series, this timeliness does *not* translate into an incremental forecasting ability for the fast-moving (high-frequency) movements in the account series. This misconception would not have been detected through conventional GC tests, and nicely illustrates the additional insights that can be obtained through the proposed spectral decomposition approach.

[Insert Table 1 about here]

The differential intensity of the GC relationship across different time horizons is also evident in Table 1, where we present the outcomes of formal spectral-based tests at three different frequencies: $\lambda_1 = 0.5$, $\lambda_2 = 1.5$ and $\lambda_3 = 2.5$. In line with our earlier discussion, we find (i) that considerable more GC is found at the low frequency (seven countries) than at the high frequency (three countries), and (ii) that for some countries (e.g. Ireland), there is little support for GC at any frequency, as opposed to countries like Belgium, France, Germany and Italy (among others) where there is significant GC at various frequencies. This differential behavior is also reflected in the final column of Table 1, which jointly considers the significance of the GC relationship across all three frequencies.

The spectral-based GC test therefore complements and extends the insights obtained through traditional CG tests. For example, from Table 1, we can infer that the highly significant overall GC tests for Finland, which are in line with earlier findings by Bergström (1995) and Teräsvirta (1986), are primarily due to the ability of the Finnish production expectation series to better predict the long-run evolution of their production accounts' counterparts. For Luxembourg, in contrast, this is due to frequencies in the 1.5-2.5 range. Also when comparing Germany and France, interesting new insights emerge. The Granger-Wald and Haugh-Pierce test both indicate, in line with Lemmens *et al.* (2004), a highly significant predictive power for their respective production expectation series. The spectral tests, however, show that for France, superior predictions are only obtained for time horizons beyond three months, while German production expectation series have significance predictive content at high frequencies as well.

4 Conclusions

As indicated before, GC has been used extensively in previous work to study a wide range of substantive economic issues. Even though there is increasing evidence that the nature of the relationship may vary with the time horizon under consideration (see e.g. Baxter, 1994), most previous studies have applied an overall GC test. In this paper, we presented a new, spectral-based, approach that, in contrast to these existing tests, offers insights into potential variations in the strength of the GC over different time horizons. We demonstrated the additional insights that could be obtained with this testing procedure in the context of the forecasting ability of European-wide expectation surveys. We believe that comparable additional insights could also be obtained in several other substantive areas.

However, a number of interesting areas for future research remain. First, while we have shown how a joint test for GC can be carried out for a finite number of distinct frequencies, we did not yet develop a test procedure to test for the nullity of $h_{u \Rightarrow v}(\lambda)$ over a subinterval $\lambda \in [a, b]$ of $[-\pi, \pi]$. The distribution of such a test statistic would be very complicated, and is beyond the scope of the current paper. As another limitation, we restricted ourselves to a test for bivariate GC. In some settings however, see e.g. Lemmens *et al.* (2004), it could be insightful to also carry out multivariate tests of GC.

References

[1] Abual-Foul, B. (2004). 'Testing the Export-Led Growth Hypothesis: Evidence from Jordan', *Applied Eonomics Letters*, Vol. 11, pp. 393-396.

[2] Al-Yousif, Y.K. (2002). 'Defense Spending and Economic Growth: Some Empirical Evidence from the Arab Gulf region', *Defence Peace Economics*, Vol. 13, pp. 187-197.

[3] Bahmani-Oskooee, M., and Niroomand, F. (1999). 'Openness and Economic Growth: An Empirical Investigation', *Applied Eonomics Letters*, Vol. 6, pp. 557-561.

[4] Barksdale, H.C., Hilliard, J.E., and Guffey, H.J. (1974). 'Spectral Analysis of Adverising-Sales Interaction: An Illustration', *Journal of Advertising*, Vol. 3, pp. 26-33.

[5] Baxter, M. (1994). 'Real Exchange Rates and Real Interest Differentials: Have We Missed the Business-Cycle Relationship?', *Journal of Monetary Economics*, Vol. 33, pp. 5-38.

[6] Bergström, R. (1995). 'The Relationship between Manufacturing Production and Different Business Survey Series in Sweden 1968-1992', *International Journal of Forecasting*, Vol. 11, pp. 379-393.

[7] Brillinger, D.R. (1969). 'Asymptotic Properties of Spectral Estimates of Second Order', *Biometrika*, 56, pp. 375-390.

[8] Bronnenberg, B.J., Mela, C.F., and Boulding, W. (2004). 'The Periodicity of Competitor Pricing', *Working paper, The Andersen School of Management, UCLA.*

[9] Bruneau, C., and Jondeau, E. (1999). 'Long-Run Causality, with an Application to International Links between Long-Term Interest Rates', *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 545-568.

[10] Buffeteau, S., and Mora, V. (2000). 'Predicting the National Accounts of the Euro Zone using Business Surveys, *INSEE Conjuncture in France*, December.

[11] Bult, J.R., Leeflang, P.S.H., and, Wittink, D.R. (1997). 'The Relative Performance of Bivariate Causality Tests in Small Samples', *European Journal of Operational Research*, Vol. 97, pp. 450-464.

[12] Chatfield, C. (1996). *The Analysis of Time Series: An Introduction*, fifth edition, London: Chapman and Hall.

[13] Cheung, Y.-W., and Fujii, E. (2001). 'A Note on the Power of Money-Output Causality Tests', *Oxford Bulletin of Economics and Statistics*, Vol. 63, pp. 247-261.

[14] Croux C., Forni, M., and Reichlin, L. (2001). 'A Measure of Comovement for Economic Variables: Theory and Empirics', *The Review of Economics and Statistics*, Vol. 83, pp. 232-241.

[15] Deleersnyder, B., Dekimpe, M.G., and Leeflang, P.S.H. (2003). 'Advertising and the Macro-Economy: Have we Missed the Business-Cycle Relationship?', *Paper presented at the 33th EMAC Conference, Murcia 2004*.

[16] Diebold, F.X. (2001). *Elements of Forecasting*, second edition, Ohio: South-Western.

[17] Dopke, J. (2004). 'How Robust is the Empirical Link between Business-Cycle Volatility and Long-Run Growth in OECD Countries?', *International Review of Applied Economics*, Vol. 18, pp. 103-121.

[18] Gouriéroux, Ch., Monfort, A., and Gallo, G. (1996). *Time Series and Dynamic Models*, Glasgow: Cambridge University Press.

[19] Granger, C.W.J. (1969). 'Investigating Causal Relations by Econometric Models and Cross-Spectral Methods', *Econometrica*, Vol. 37, pp. 424-438.

[20] Hanssens, D.M., Parsons, L.J., and Schultz, R.L. (2001). *Market Response Models: Econometric and Time Series Analysis*, 2nd Edition, Kluwer Academic Publishers.

[21] Hanssens, D.M., and Vanden Abeele, P.M. (1987). 'A Time-Series Study of the Formation and Predictive Performance of EEC Production Survey Expectations', *Journal of Business and Economic Statistics*, Vol. 5, pp. 507-519.

[22] Hemstra, C., and Jones, J.D (1994). 'Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation', *Journal of Finance*, Vol. 49, pp. 1639-1664.

[23] Koopmans, L.H. (1995). *The Spectral Analysis of Time Series*, Probability and Mathematical Statistics 22, San Diego: Academic Press.

[24] Lemmens, A., Croux, Ch., and Dekimpe, M.G. (2004). 'The Predictive Content of Production Surveys: a Pan-European Study', *International Journal of Forecasting*, forthcoming.

[25] Öller, L.-E. and Tallbom, C. (1996). 'Smooth and Timely Business Cycle Indicators for Noisy Swedish Data', *International Journal of Forecasting*, 12, pp. 389-402.

[26] Park, J., and Shenoy, C. (2002). 'An Examination of the Dynamic Behavior of Aggregate Bond and Stock Issues', *International Review of Economics and Finance*, Vol. 11, pp. 175-189.

[27] Peiers, B. (1997). 'Informed Traders, Intervention, and Price Leadership: A Deeper View of the Microstructure of the Foreign Exchange Market', *Journal of Finance*, Vol. 52, pp. 1589-1614.

[28] Pierce, D.A., and Haugh, L.D. (1977). 'Causality in Temporal Systems', *Journal of Econometrics*, Vol. 5, pp. 265-293.

[29] Teräsvirta, T. (1986). 'Model Selection with Business Survey Data: Forecasting the Output of Finnish Metal and Engineering Industries', *International Journal of Forecasting*, Vol. 2, pp. 191-200.

[30] Woitek, U. (2003). 'Height Cycles in the 18th and 19th Centuries', *Economics and Human Biology*, Vol. 1, pp. 243-257.

[31] Zellner, A. (1979). 'Causality and Econometrics', *Journal of Monetary Economics (Supplement)*, Vol. 10, pp. 1-54.

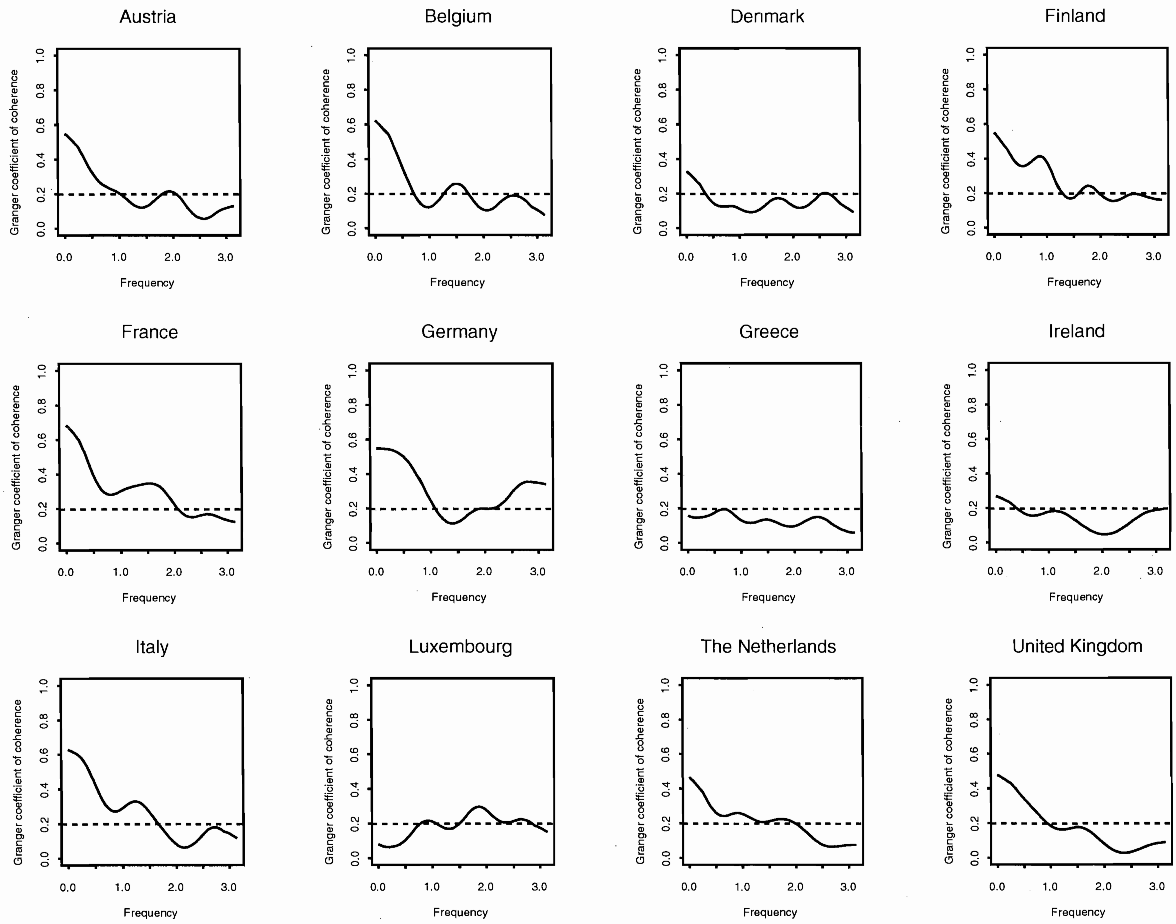

Figure 1: Granger coefficients of coherence for 12 European countries. The dotted line represents the critical bound at the 5% probability level.

p-value	Traditional tests		Spectral-based tests			
Country	G-W test	H-P test	$\lambda_1 = 0.5$	$\lambda_2 = 1.5$	$\lambda_3 = 2.5$	Joint test ($s = 3$)
Austria	**0.006**	**0.008**	**0.012**	0.400	0.946	0.069
Belgium	**0.000**	**0.003**	**0.005**	**0.007**	0.065	**0.000**
Finland	**0.000**	**0.000**	**0.007**	0.221	0.052	**0.006**
France	**0.000**	**0.000**	**0.003**	**0.001**	0.094	**0.000**
Germany	**0.000**	**0.001**	**0.000**	0.398	**0.004**	**0.000**
Italy	**0.000**	**0.001**	**0.000**	**0.008**	0.152	**0.000**
The Netherlands	0.056	0.057	0.071	0.054	0.616	0.065
United Kingdom	**0.002**	0.132	**0.003**	0.077	0.790	**0.008**
Denmark	0.146	0.193	0.453	0.196	**0.038**	0.089
Greece	**0.003**	0.080	0.067	0.146	0.120	**0.048**
Ireland	0.357	0.213	0.160	0.208	0.292	0.197
Luxembourg	**0.001**	**0.013**	0.284	0.130	**0.038**	0.051

Table 1: p-values for the Granger-Wald (G-W) test, Haugh-Pierce (H-P) test, and the spectral-based GC test at three different frequencies for testing GC between production expectations and accounts (with $M = \sqrt{T}$). Significant p-values at the 5% probability level are reported in bold.